WARSAW DOCTORAL SCHOOL OF MATHEMATICS AND COMPUTER SCIENCE

July 5, 2023

ENTRANCE EXAM

On the following pages you will find 16 problems related to various areas of Mathematics and Computer Science. You are expected to choose and solve **any 4 of them**. Each problem is worth the same number of points.

You are free to choose any problems you wish, i.e. candidates for studies in Mathematics may also choose Computer Science problems, and *vice versa*.

Most of the problems are composed of a few subproblems, but each problem, i.e. all of its subproblems, are graded as a whole.

You may attempt to solve more than 4 problems. All your solutions will be graded, but **only 4 best-graded solutions will contribute to your general grade**.

All your answers should be appropriately justified. **Every problem should be solved on a separate sheet of paper**; of course, the solution of one problem can be written on more that one sheet.

Each sheet of paper should be **signed with your first name and surname**, and **marked with the problem number**.

EXAM DURATION: 3 HOURS

Good luck!

Analysis

PROBLEM 1. Let $I \subset \mathbb{R}$ be an open interval. We say that a function $f\colon I \to \mathbb{R}$ is convex if $f(tx + (1-t)y) \leqslant tf(x) + (1-t)f(y)$ for all $x, y \in I$, $t \in [0,1]$.

(a) Prove that if $\varphi\colon I \to \mathbb{R}$ is a convex function, then for every $x \in I$ and every $h > 0$ with $x-h, x+h \in I$, the following inequality holds true:

$$\mathcal{I}(x,h) := \frac{1}{2h}\int_{-h}^{h} \varphi(x+t)\,\mathrm{d}t \geqslant \varphi(x).$$

(b) Suppose that $\varphi\colon (0,\infty) \to \mathbb{R}$ is a continuous function for which the above-defined integral $\mathcal{I}(x,h)$ has constant value for all $x, h \in (0,\infty)$ such that $h < x$. Does it imply that φ is a constant function?

(c) Assume that a function $f\colon \mathbb{R}^2 \to \mathbb{R}$ is differentiable on $\mathbb{R}^2 \setminus \{(0,0)\}$ and satisfies the following conditions:

$$\begin{cases} yf_x(x,y) - xf_y(x,y) = 0 & \text{for } (x,y) \neq (0,0) \\ f(x,0) = x^2e^{-x} & \text{for } x \geqslant 0. \end{cases}$$

Derive an explicit formula for f and prove that f has exactly one strict local minimum and infinitely many non-strict local maxima.

(d) Find the largest $k \in \mathbb{N}$ for which f considered in point (c) is k-times differentiable on $\mathbb{R}^2$, or prove that f is differentiable infinitely many times on $\mathbb{R}^2$.

Complex analysis

PROBLEM 2. Let $\Omega \subset \mathbb{C}$ be a nonempty bounded connected open set, and let $f, g\colon \overline{\Omega} \to \mathbb{C}$ be continuous on the closure $\overline{\Omega}$ and holomorphic in Ω. By $\partial\,\Omega$ we denote the boundary of Ω.

(a) Prove that

$$\max_{z\in\overline{\Omega}}(|f(z)| + |g(z)|) = \max_{z\in\partial\,\Omega}(|f(z)| + |g(z)|).$$

(b) Let $\Omega = D = \{|z| < 1\}$ be the open unit disc and let $f, g\colon \overline{D} \to \mathbb{C}$ be as above. Prove that the following conditions:

- $|f(z)| \leqslant 1$, $|g(z)| \leqslant 2$ for $|z| = 1$, $\operatorname{Im} z \geqslant 0$,
- $|f(z)| \leqslant 3$, $|g(z)| \leqslant 1$ for $|z| = 1$, $\operatorname{Im} z \leqslant 0$,

imply that $|f(0)| \cdot |g(0)| \leqslant \frac{5}{2}$.

(c) Now, assume that Ω is a region containing the closure $\overline{D}$ and $f, g\colon \Omega \to \mathbb{C}$ are holomorphic functions such that:

- $|f(z) + z| < 1$ for $|z| = 1$,
- $|g(z)| < 1$ for $|z| < 1$.

Show that the function

$$h(z) := f(z) + \frac{g(z) - g(0)}{1 - \overline{g(0)}g(z)}$$

has at least one fixed point, i.e. there is $z \in \Omega$ with $h(z) = z$.

(d) Moreover, show that under the assumptions in (c), the function h has exactly one fixed point in the disc D, provided that $g'(0) = 0$.

Probability and statistics

Problem 3. Let $\xi_1, \xi_2, \ldots$ be a Bernoulli sequence with parameter $x \in (0,1)$, that is, a sequence of independent random variables with $\mathbb{P}(\xi_k = 1) = x$, $\mathbb{P}(\xi_k = 0) = 1 - x$ for each $k \in \mathbb{N}$. Let X be the random variable that counts failures until the first occurrence of success in the Bernoulli process, i.e. $X = \min\{k \in \mathbb{N} \colon \xi_k = 1\} - 1$.

(a) Find the distribution of X, and calculate the expected value $\mathbb{E}X$ and the variance $\mathbb{D}^2 X$.

(b) Let $X_1, X_2, \ldots$ be a sequence of independent copies of X, and for any $n \in \mathbb{N}$, we denote $S_n = X_1 + \ldots + X_n$. Show that for each $k = 0, 1, 2, \ldots$, we have

$$\mathbb{P}(S_n = k) = \binom{n+k-1}{k}(1-x)^k x^n.$$

(c) Calculate the limit

$$\lim_{n\to\infty} x^n \sum_{k=0}^{\lfloor (1-x)n/x \rfloor} \binom{n+k-1}{k}(1-x)^k,$$

where $\lfloor t \rfloor$ stands for the integer part of t.

(d) Let $\{0,1\}^{\mathbb{N}}$ be equipped with the product measure, where on which copy of $\{0,1\}$ we take the uniform measure $\mu(\{0\}) = \mu(\{1\}) = \frac{1}{2}$. For any $\overline{\omega} = (\omega_n)_{n=1}^{\infty} \in \{0,1\}^{\mathbb{N}}$, and any $k \in \mathbb{N}$, define $\mathsf{i}_k(\overline{\omega}) \in \mathbb{N} \cup \{\infty\}$ as the index of k^{th} occurrence of '1' in the sequence $\overline{\omega}$, and we set $\mathsf{i}_k(\overline{\omega}) = \infty$ if there are less than k appearances of '1' in $\overline{\omega}$. Let also $\mathsf{i}_0(\overline{\omega}) \equiv 0$.

- Let $A \subset \{0,1\}^{\mathbb{N}}$ be a set of positive measure. Show that there exist $\overline{\omega} \in A$ and $n_0 \in \mathbb{N}$ such that

$$(\mathsf{i}_1(\overline{\omega}) - \mathsf{i}_0(\overline{\omega}))^2 + (\mathsf{i}_2(\overline{\omega}) - \mathsf{i}_1(\overline{\omega}))^2 + \ldots + (\mathsf{i}_n(\overline{\omega}) - \mathsf{i}_{n-1}(\overline{\omega}))^2 < 7n$$

  for all $n \geqslant n_0$.

- Prove that $\mathbb{P}\Big(\mathsf{i}_k + \mathsf{i}_{k-2} > 2\mathsf{i}_{k-1} \text{ for each } k \geqslant 2\Big) = 0$.

Geometry and linear algebra

Problem 4. Let V be a finite-dimensional vector space over $\mathbb{C}$, the field of complex numbers, with $n = \dim V \geqslant 2$. We define a $k \times k$ *Jordan block* with parameter $\lambda \in \mathbb{C}$ as

$$\mathbf{J}_{\lambda,k} = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & \lambda \end{pmatrix}$$

(a) Let $\lambda \neq 0$, $1 \leqslant k \leqslant n$, and $\mathbf{A} \in \mathbb{M}_{n\times n}(\mathbb{C})$ be the block matrix $\mathbf{A} = \mathbf{J}_{0,k} \oplus \mathbf{J}_{\lambda,n-k}$, that is,

$$\mathbf{A} = \left(\begin{array}{c|c} \mathbf{J}_{0,k} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{J}_{\lambda,n-k} \end{array}\right).$$

Determine the numbers:

- $\nu := \max\{\dim\ker \mathbf{A}^m \colon m \in \mathbb{N}\}$,
- $\mu := \min\{m \in \mathbb{N} \colon \dim\ker \mathbf{A}^m = \nu\}$.

(b) Suppose $T\colon V \to V$ is a linear endomorphism of V such that

$$(\star) \qquad \ker T^{n-2} \neq \ker T^{n-1}.$$

Prove that T has at most two distinct eigenvalues.

(c) Show that if T satisfies $(\star)$ and T is diagonalizable, then $n = 2$.

(d) Assuming still that T satisfies $(\star)$, let $S\colon V \to V$ be a linear endomorphism such that the operators $T + S$ and $T - S$ commute and are Hermitian (self-adjoint). Prove that $n = 2$.

Algebra

Problem 5. Throughout this problem we assume that G is a finite group of odd order $n = |G|$. It is known, and not difficult to prove, that the square function $G \ni x \mapsto x^2$ is then injective, and that this property is in fact equivalent to the order of G being odd.

(a) Show that for each $x \in G$ there is an integer $k \geqslant 1$ such that $x^{2^k} = x$.

(b) Let N be a normal subgroup of G and $a, b \in G$. Show that if $a^2 b^{-2} \in N$, then $ab^{-1} \in N$.

(c) In view of assertion (a), there is a correctly defined function $\ell\colon G \to \{1, \ldots, n\}$ such that $\ell(x)$ is the smallest natural number k for which $x^{2^k} = x$. Prove that if ℓ is injective, and G has at least three elements, then n is divisible by 21.

(d) Let $\mathbb{C}[G]$ stand for the group ring of G over the complex numbers, that is, the ring consisting of all functions $\phi\colon G \to \mathbb{C}$ with pointwise addition, i.e. $(\phi + \psi)(s) = \phi(s) + \psi(s)$ for $\phi, \psi \in \mathbb{C}[G]$ and $s \in G$, and multiplication given as the convolution

$$(\phi \cdot \psi)(s) = \sum_{t \in G} \phi(t)\psi(t^{-1}s).$$

Notice that this ring is noncommutative if G is noncommutative, and that every element ϕ of $\mathbb{C}[G]$ can be written as a formal combination $\phi = \sum_{s \in G} a_s s$, where $a_s = \phi(s)$, and then the multiplication between two such combinations is just an ordinary term-by-term multiplication.

Suppose $g \in G$ is an element of order 3 and, using the above convention, let $\mathscr{I} \subseteq \mathbb{C}[G]$ be the right ideal generated by $\phi = 1 - g$, where 1 is the neutral element of G. Define also

$$\mathscr{M} = \Big\{ \sum_{s \in G} a_s s \in \mathbb{C}[G] \colon \sum_{s \in G} a_s = 0 \Big\}.$$

- Show that $1 - g$ is a zero divisor in $\mathbb{C}[G]$.
- Show that if $h \in G \setminus \{1, g, g^2\}$, then $1 - h \notin \mathscr{I}$.
- Prove that $\mathscr{I} \subseteq \mathscr{M}$ and that $\mathscr{M}$ is a maximal proper two-sided ideal of $\mathbb{C}[G]$.
- Is the quotient ring $\mathbb{C}[G]/\mathscr{M}$ commutative?

Topology

Problem 6. Consider the family $\mathcal{A}$ consisting of the empty set and all arithmetic progressions in $\mathbb{Z}$, i.e. sets of the form

$$S(a, b) = \{an + b \colon n \in \mathbb{Z}\},$$

where $a \neq 0$.

(a) Show that the collection of all unions of elements of $\mathcal{A}$ constitutes a topology $\mathcal{T}$ on $\mathbb{Z}$ with a basis $\mathcal{A}$.

(b) Verify whether $\mathcal{T}$ satisfies the separation axioms: T_0, T_1 and T_2.

(c) Determine the general form of finite open sets in $\mathcal{T}$. Is $\mathbb{Z}$ a discrete space in the topology $\mathcal{T}$?

(d) Show that the sets from the basis $\mathcal{A}$ are clopen (both closed and open).

(e) Decide if the set of all numbers that are not a multiple of any prime number, i.e. the set $\{-1, 1\}$, is open in $\mathcal{T}$. Using assertion (d) conclude that there are infinitely many primes.

Ordinary differential equations

PROBLEM 7. Define a function $f\colon \mathbb{R} \to \mathbb{R}$ by the formulas

$$f(x) = \begin{cases} (x+1)^{2/3} & \text{for } x \leqslant 0 \\ 2\exp(-x^2) - 1 & \text{for } 0 < x < 1 \\ \dfrac{\mathrm{e}-2}{3\mathrm{e}}(x^2-4) & \text{for } x \geqslant 1. \end{cases}$$

We consider the following differential equation:

$$(\spadesuit) \qquad \frac{\mathrm{d}x}{\mathrm{d}t} = f(x).$$

(a) Determine all solutions of $(\spadesuit)$ satisfying the initial condition $x(0) = -1$ that are defined on some neighborhood of $t = 0$.

(b) Let x_b be the solution of equation $(\spadesuit)$ with an initial condition $x(0) = x_0 > -1$. Determine the maximal interval (ω_0, ω_+) on which x_b can be defined, depending on the parameter x_0.

(c) Decide whether the left-side limit exists:

$$\lim_{t \to \omega_+^-} x(t),$$

and if it does, calculate it. (The answer may depend on x_0.)

(d) Assume that $g_0\colon \mathbb{R} \to \mathbb{R}$ is a continuous function satisfying $g_0(t) = g_0(t+1)$. Prove that there is a unique constant $a \in \mathbb{R}$ such that the solution to the Cauchy problem

$$\frac{\mathrm{d}y}{\mathrm{d}t} + 2y = g_0(t), \quad y(0) = a$$

is a periodic function of period 1. Denote that solution by y^*.

(e) Show that for every continuous function $g\colon \mathbb{R} \to \mathbb{R}$ satisfying $|g(t) - g_0(t)| < e^{-t}$ for each $t > 0$, every solution to the Cauchy problem

$$\frac{\mathrm{d}y}{\mathrm{d}t} + 2y = g(t), \quad y(0) = y_0$$

satisfies $\lim_{t \to +\infty} (y(t) - y^*(t)) = 0$.

Functional analysis

PROBLEM 8. Let $L^2(\mathbb{R})$ stand for the classical Hilbert space of (classes of abstraction of) measurable functions $f\colon \mathbb{R} \to \mathbb{C}$ such that $\|f\|_2 \coloneqq \{\int_{\mathbb{R}} |f(x)|^2 \,\mathrm{d}x\}^{1/2} < \infty$. Let also $C_0(\mathbb{R})$ be the Banach space of bounded continuous functions $f\colon \mathbb{R} \to \mathbb{C}$ such that $\lim_{x \to -\infty} f(x) = 0$, equipped with the supremum norm, $\|f\|_\infty \coloneqq \sup_{x \in \mathbb{R}} |f(x)|$. Fix an arbitrary $g \in L^2(\mathbb{R})$ and define a linear operator on $L^2(\mathbb{R})$ by the formula

$$(\clubsuit) \qquad Tf(t) = \int_{-\infty}^{t} f(x) g(x) \,\mathrm{d}x.$$

(a) Show that for each $f \in L^2(\mathbb{R})$ the above-defined function Tf belongs to $C_0(\mathbb{R})$, hence formula $(\clubsuit)$ defines a linear operator $T\colon L^2(\mathbb{R}) \to C_0(\mathbb{R})$. Calculate $\|T\|$.

(b) Let $C[0,1]$ be the Banach space of continuous complex-valued functions defined on the interval $[0,1]$, equipped with the supremum norm. Consider the operator $S\colon C_0(\mathbb{R}) \to C[0,1]$ given as $Sf = f\restriction_{[0,1]}$. Show that the composition $ST\colon L^2(\mathbb{R}) \to C[0,1]$ is a compact operator.

(c) Now, assume that $g = \mathbb{1}_{[0,1]}$. Consider the operator $\iota\colon C[0,1] \to L^2(\mathbb{R})$ given by $\iota f(t) = f(t)$ for $t \in [0,1]$, and $\iota f(t) = 0$ for $t \in \mathbb{R} \setminus [0,1]$, and define $U\colon L^2(\mathbb{R}) \to L^2(\mathbb{R})$ by $U = \iota ST$. Prove that the spectrum of this operator $\sigma(U) = \{0\}$.

(d) Assuming again that $g = \mathbb{1}_{[0,1]}$ determine the adjoint operator U^* of U and calculate the norm $\|U^*U\|$.

Programming languages

PROBLEM 9.

(a) Provide the result of the following function `f` in the C language, when called with a pointer to an array {`0, 0, 0, 0, 0, 1, -1`}:

```
int f(int *A) {
    int tmp = *A;
    tmp += tmp < 0 ? -tmp : f(++A);
    if (*(A++) >= 0) tmp += f(A);
    return tmp % 1000;
}
```

Write a function computing the same as `f`, but without using recursive calls (and without allocating additional arrays).

(b) Write functions `f` and `g` in the OCaml language, such that the expressions

```
List.fold_right f [(1, 2); (3, 4); (5, 6)] ([], [])
List.fold_right f [('a', 'b')] ([], [])
List.map g [8; 9; 11; 12]
```

return, respectively,

```
([1; 3; 5], [2; 4; 6])
(['a'], ['b'])
[12; 11; 9; 8]
```

(c) Consider the following function in the C language:

```
void f(char *S) {
    for (int a = 0; a < strlen(S); ++a)
        if (S[a] >= 'a' && S[a] <= 'z')
            S[a] = 'z' - (S[a] - 'a');
}
```

The function is unexpectedly slow: tested on some modern computer, it worked around 14 seconds for a string of length 10^6. What is the reason? How to correct this function?

(d) Consider a similar function in the Java language:

```
String f(String S) {
    String res = "";
    for (int a = 0; a < S.length(); ++a) {
        char ch = S.charAt(a);
        res += (ch >= 'a' && ch <= 'z') ? (char)('z' - (ch - 'a')) : ch;
    }
    return res;
}
```

Also this function is very slow: it worked around 57 seconds for a string of length 10^6. What is the reason? How to correct this function?

Discrete mathematics

PROBLEM 10. Given a positive integer n, a *permutation* of $[n] := \{1, 2, \dots, n\}$ is a reordering of the numbers $1, \dots, n$, that is, a bijective function $\pi : [n] \to [n]$. Thus $(3, 1, 2)$ is a permutation of $\{1, 2, 3\}$, with $\pi(1) = 3, \pi(2) = 1, \pi(3) = 2$.

(a) Given $1 \leqslant i < j < k \leqslant n$, how many permutations π of $[n]$ are there such that $\pi(j) < \pi(i) < \pi(k)$?

(b) Given $j \in [n]$, how many permutations π of $[n]$ are there such that for all $i \in [j]$ we have $\pi(i) \leqslant \pi(j)$?

(c) Given $j \in [n]$ and $i \in [j]$, how many permutations π of $[n]$ are there, such that there are exactly i positions in $[j]$, mapped to numbers not greater than $\pi(j)$?

(d) Given numbers $1 \leqslant i_1 < i_2 < \ldots < i_r \leqslant n$, how many permutations π of $[n]$ are there, such that for all $k \in [r]$ and $j \in [i_k]$ we have $\pi(j) \leqslant \pi(i_k)$?

For each of the above cases, you should provide a concise formula (in terms of the given quantities).

Algorithms and data structures

PROBLEM 11. You are given an array A of length n containing natural numbers. For each of the following problems, propose an algorithm working in linear time and with constant additional memory. Be precise: write a pseudo-code and explain non-obvious details; justify correctness and complexity.

(a) Find three indices $i < j < k$ such that $A[i] < A[j] < A[k]$ (or say that they do not exist).

(b) Compute the number of pairs $i < j$ such that $A[i] < A[i + 1] < \cdots < A[j]$.

(c) Given a natural number d and assuming that the array is sorted, find the largest k such that for some i we have $A[k + i] - A[i] = d$ (or say that it does not exist).

(d) Given a natural number s find the largest k such that for some i we have $A[i] + A[i+1] + \cdots + A[i + k] = s$ (or say that it does not exist).

Logic and databases

PROBLEM 12. Consider the first-order logic with equality and with one binary relational symbol $\leqslant$. Let ψ_{lin} be a sentence saying that $\leqslant$ is a linear order. Does there exist a sentence φ such that

(a) $\psi_{lin} \wedge \varphi$ has models, but no finite models;

(b) $\psi_{lin} \wedge \varphi$ has arbitrarily large finite models, but only of even size;

(c) all countable models of $\psi_{lin} \wedge \varphi$ are isomorphic to the set of natural numbers;

(d) all countable models of $\psi_{lin} \wedge \varphi$ are isomorphic to the set of rational numbers?

Two models are isomorphic if there is a bijection f between them satisfying $x \leqslant y \Leftrightarrow f(x) \leqslant f(y)$.

Automata and formal languages

PROBLEM 13. For a binary sequence $w \in \{0, 1\}^*$ let $[w]_2$ denote its numerical value; e.g., $[011]_2 = 3$. Below, we consider alphabets $\{0, 1\}^k$ (where $k = 4$ in L_1; $k = 2$ in L_2 and in L_3; and $k = 3$ in L_4), with letters written as columns. For each of the following languages, say (and prove) whether or not it is regular and whether or not it is context-free.

$$L_1 = \left\{ \begin{bmatrix} a_1 \\ b_1 \\ c_1 \\ d_1 \end{bmatrix} \begin{bmatrix} a_2 \\ b_2 \\ c_2 \\ d_2 \end{bmatrix} \dots \begin{bmatrix} a_n \\ b_n \\ c_n \\ d_n \end{bmatrix} : [a_1 a_2 \dots a_n]_2 + [b_1 b_2 \dots a_n]_2 = [c_1 c_2 \dots c_n]_2 + [d_1 d_2 \dots d_n]_2 \right\}$$

$$L_2 = \left\{ \begin{bmatrix} a_1 \\ b_1 \end{bmatrix} \begin{bmatrix} a_2 \\ b_2 \end{bmatrix} \dots \begin{bmatrix} a_n \\ b_n \end{bmatrix} \begin{bmatrix} c_1 \\ d_1 \end{bmatrix} \begin{bmatrix} c_2 \\ d_2 \end{bmatrix} \dots \begin{bmatrix} c_n \\ d_n \end{bmatrix} : \begin{aligned} &[a_1 a_2 \dots a_n]_2 + [b_1 b_2 \dots a_n]_2 = \\ &\qquad [c_1 c_2 \dots c_n]_2 + [d_1 d_2 \dots d_n]_2 \end{aligned} \right\}$$

$$L_3 = \left\{ \begin{bmatrix} a_1 \\ b_1 \end{bmatrix} \begin{bmatrix} a_2 \\ b_2 \end{bmatrix} \dots \begin{bmatrix} a_n \\ b_n \end{bmatrix} \begin{bmatrix} c_n \\ d_n \end{bmatrix} \begin{bmatrix} c_{n-1} \\ d_{n-1} \end{bmatrix} \dots \begin{bmatrix} c_1 \\ d_1 \end{bmatrix} : \begin{aligned} &[a_1 a_2 \dots a_n]_2 + [b_1 b_2 \dots a_n]_2 = \\ &\qquad [c_1 c_2 \dots c_n]_2 + [d_1 d_2 \dots d_n]_2 \end{aligned} \right\}$$

$$L_4 = \left\{ \begin{bmatrix} a_1 \\ b_1 \\ c_1 \end{bmatrix} \begin{bmatrix} a_2 \\ b_2 \\ c_2 \end{bmatrix} \dots \begin{bmatrix} a_n \\ b_n \\ c_n \end{bmatrix} : [a_1 a_2 \dots a_n]_2 \cdot [b_1 b_2 \dots a_n]_2 = [c_1 c_2 \dots c_n]_2 \right\}$$

Computation theory and computational complexity

PROBLEM 14. Determine the complexity of each of the following problems—choose from: in PTIME, NP-complete, PSPACE-complete, EXPTIME-complete, EXPSPACE-complete, undecidable. Provide proofs for your claims.

(a) **Input:** Nondeterministic single-tape Turing machine M.
Question: Is there an input word on which all runs of M halt in at most 2023 steps?

(b) **Input:** Deterministic single-tape Turing machine M.
Question: Is there an input word w on which M halts in at most $\log_{2023} |w|$ steps?

(c) **Input:** Deterministic single-tape Turing machine M; number n (written in binary).
Question: Is there an input word of length n on which M halts in at most $\log_{2023} n$ steps?

(d) **Input:** Deterministic single-tape Turing machine M; number n (written in binary).
Question: Is there an input word w such that the (finite or infinite) run of M on w has the following property: for every $k \in \mathbb{N}$, if the run visits at some moment cell number k, then later it never visits cell number $k - \lfloor \sqrt[2023]{n} \rfloor$?

Concurrent and distributed programming, computer systems

PROBLEM 15. *Fork-join* is a parallel programming model for shared memory. As an illustration, consider the following code listing in the C programming language with enhancements for supporting the model represented as underlined keywords:

```
int fibonacci(int i) {
  if (i <= 1) {
    return 1;
  } else {
    task t = fork fibonacci(i - 1);
    int f2 = fibonacci(i - 2);
    int f1 = join t;
    return f1 + f2;
  }
}
```

The **fork** instruction in line 5 creates a new (child) **task**, that is, an invocation of function `fibonacci(int)` with a specific parameter value `(i − 1)`. After creation, this task is to be executed asynchronously: depending on the available processors, it can proceed in parallel to the (parent) task that forked it, and notably, to the recursive call of `fibonacci(i − 2)` from line 6 of the parent task. The **join** instruction in line 7 causes the invoking parent task to wait for the forked child task to complete (if it has not completed already) and collects its result. In effect, the function from the listing computes (in a naïve way) the i-th Fibonacci number, employing potentially many processors to this end.

In general, the model assumes that the number of tasks forked at a given moment can be unlimited. In particular, before waiting for a child task to complete, a parent task can fork other child tasks. However, before completing, any task has to wait for the completion of all its child tasks. On the other hand, no task can wait for the completion of another task unless it is the parent of that task. If there are sufficiently many processors available, forked tasks can all execute in parallel; otherwise, some of them have to wait for an available processor. The scheduling of the tasks onto processors is transparent to the programmer.

Different tasks can access shared memory (e.g., global variables or arrays passed as function parameters). Such accesses are realized by two atomic operations, load and store, which operate only on values of type **int**. If multiple tasks access the same memory location (e.g., the same array element) at the same time, the load and store operations corresponding to these accesses are arbitrated. More specifically, while multiple loads can take place at the same time, stores are mutually exclusive with each other (i.e., they are in fact performed in sequence) and with loads. The order of such mutually exclusive operations is beyond the control of the programmer. Load and store are the only supported atomic operations.

Your objective is to write an efficient parallel implementation of a function for performing prefix sums in the fork-join model. Given an n-element input array x of integer values, prefix sums of x are an n-element array y such that:

$$y[i] = \begin{cases} x[0] & \text{if } i = 0, \\ y[i-1] + x[i] & \text{otherwise.} \end{cases}$$

More specifically, you are to implement the following function:

```
void prefix_sums(int const * x, int * y, int n);
```

where `x` is the input array with the integer values, `y` is the output array that should contain the prefix sums after the function completes, and `n` is the length of both arrays. You have to provide the complete code for your solution (a mere description will be graded very low). In particular, if you use helper functions, you should implement them all as well. Add comments on nontrivial aspects of your implementation. Besides the implementation, you should provide asymptotic bounds for its work (the total number of operations executed by all processors) and span (the running time assuming arbitrarily many processors). You should include not only the final formulae but also their derivations. If your solution utilizes any extra arrays, you can assume that the asymptotic cost of allocating an uninitialized array or freeing one is $\Theta(1)$. If you need to prioritize between work and span, give preference to a smaller span.

Bioinformatics

PROBLEM 16.

(a) A short-read sequencing experiment resulted in the following collection of DNA reads: {ACGTGT, GTCATT, ATTACG, GTGTCA}. Perform *de novo* assembly of the input sequence using a de Bruijn graph.

   (i) Build a de Bruijn graph for k-mer length $k = 3$.

(ii) Assemble the input sequence by finding an Eulerian path for the de Brujin graph you have constructed. Mark the path on the graph by numbering the arcs and write down the resulting sequence.

(b) What is the memory and time complexity of building the de Bruijn graph depending on the number of reads (N), length of reads (n), and the k-mer length (k)?

(c) Use the Eulerian path approach to reconstruct the query sequence knowing that its 3-mer composition is: {ACA, CAT, CAC, ACT, GCA, CTG, GGC} but one 3-mer is missing in it. What is the missing 3-mer and the query sequence?